

인터넷 환경에서 데이터 인접성에 기반한 정보 검색 및 분석을 위한 웹페이지 시각화

변현수, 김진화

서강대학교 경영학과

Visualization of web pages for information search and analysis based on data adjacency in Internet Environment

Hyeonsu Byeon, Jinhwa Kim

Abstract

As a lot of information and media are given to users in Internet space nowadays, users feel disoriented or “lost in space” intensively. So it is suggested that we have the system to reduce information overload and to propose effective and efficient information. In this study we present a visualizing technique which uses fisheye views on data adjacency to combine global context and local details for presentation of many results in limited space. Data Adjacency on graph theory is applied to set up degree of interest which is main focus in fisheye views. Graph theory is useful to solve the problem resulted from various combinational optimization, especially it has advantages to analyze issues in information space like Internet. To test the usability of the proposed visualization technique, we compared the effectiveness of different visualization techniques. Results show that our method is evaluated with respect to less time and high satisfaction for a task accomplishment.

1. 서론

인터넷의 활용이 점차 커져 감에 따라 인터넷 사용자들은 하루에도 수많은 웹 사이트를 방문하면서 여러 가지 정보 및 서비스를 활용한다. 사용자들이 접근하게 되는 개체는 그 양이 막대할 뿐만 아니라 데이터 형식 또한 텍스트, 사운드, 이미지, 동영상 등 다양하다. 이로 인해 사용자는 인터넷이라는 거대한 정보공간에서 어디에 있는지, 그리고 어떻게 왔던 곳으로 돌아가야 할지 망각하게 되는 이른바 “초공간에서의 방향 상실(lost in space)” 현상을 자주 접하게 된다(Fisher et al., 1997, Muchaluat et al., 1998, Rivlin et al., 1994). 따라서 지나치게 많은 정보과잉(information overload)의 현상 상황에서 사용자가 필요로 하는 정보를 효과적이면서 효율적으로 구현할 필요가 있다. 정보 표현의 문제해결에 있어서 요점은 정확한 정보(right information)를 필요로 하는 사람(right person)에게 적시(right time)에 올바른 형태(right form)로 전달해 주는 것이라고 할 수 있다(Turetken and Sharda, 2001).

컴퓨터 화면이라는 한정된 공간 내에서 사용자가 요구하는 결과물을 군집화하고 그것들을 표시하기 위한 시각화 기법은 정보 공간을 파악하기에 유용한 방법이다. 사용자에게 분석결과를 제시하는데 있어서 시각화 기법은 통계적 수치를 데이터의 분포로 표현함으로써 지식의 이해를 돕고 정보의 관찰을 쉽고 빠르게 하도록 할 수 있다. 이처럼 정보의 적절한 시각화는 여타 시각적 수단을 사용하여 시스템을 보조하는 유용한 도구로 쓰이며(Tufte, 2001), 전통적으로 나무 구조와 군집화 기법들이 주로 사용되었

다. 인터넷 공간의 많은 정보들은 거의 대부분 웹페이지의 형태로 존재하고 있으며, 특히 인터넷과 웹사이트의 활용이 활발한 지금 웹 사이트를 보다 효과적으로 구현하기 위해 웹 사이트 내의 콘텐츠를 보다 적확(的確)하게 표현하기 위한 방법들도 많이 개발되고 있다(Lamping and Rao, 1996; Rivlin et al., 1994).

시각화에 있어서 가장 중요하게 고려되는 부분은 전체적 맥락(context)과 상세 내용 보기(details)를 어떻게 통합시켜 제시하느냐 하는 것이다. 본 연구에서는 웹페이지에 개체들을 구현하는 시각화 기법으로서 데이터 인접성예기반을 둔 fisheye views를 이용하여 보다 시각화된 웹 구축 방법을 제시한다. Fisheye views 개념은 모든 개체를 되도록 한 화면 내에 표현하되 관심영역 부분을 부각시키는 것이며(Furnas, 1986), 여러 시각화 처리 기법 중에서도 상당히 주목받아 온 개념이다. 데이터 인접성은 그래프 이론에 기반을 두고 있으며, 다양한 조합적 문제를 모델화하고 해결하는데 유용하다. 사용자의 방문행동은 순차적 흐름을 따르게 되면서 수직적 나무 구조를 이루게 된다. 또한 이러한 과정에서 데이터 간의 관계성이 구축되는데 이러한 수평적 연관성을 함께 고려한 다음 시각화에 적용하여 사용자에게 통찰력을 주는 정보 표현을 구체화하여야 한다. 본 연구를 통해서 보다 효율적인 웹사이트 설계 방식을 제안할 수 있으리라 기대한다.

2. 관련 연구

2.1 시각화와 fisheye views

데이터 시각화란 데이터의 구조를 일목요연하게 볼 수 있는 보기형식을 강조하는 데이터 분석 방법이라고 할 수 있다(Cleveland, 1993). 복잡한 데이터셋에서 나타나는 정보 유형을 관찰할 때, 분석가에게 쉽고 빠른 분석작업을 제공함으로써 효율성을 높이는 것이 시각화의 장점이다. 따라서 이러한 시각화 기법은 데이터마이닝 등 데이터베이스를 기반으로 하는 작업에 많이 사용된다. Becker et al.(1997)은 다차원 데이터의 시각화는 전반적인 데이터베이스내의 지식탐사(KDD: Knowledge Discovery in Databases) 과정 중 일부로서, 데이터 마이닝을 위한 주요 도구 중 하나라고 하였다. 시각화 기법은 통계적 분석을 사용하지 않고도 데이터의 분포 또는 이상점(outlier)의 존재여부 등을 개괄적이고 직관적으로 이해하도록 한다. 이러한 정보 시각화 과정은 인간과 컴퓨터와의 상호작용 결과이며 크게 2단계로 나누어 생각해 볼 수 있다(Yang et al., 2003). 하나는 정보 해석 및 매핑(information interpretation and mapping)이며, 또 하나는 정보 표현 및 통제(information display and control)이다. 먼저 데이터가 가지고 있는 다차원의 정량적이고 정성적인 특징들을 2차원 또는 3차원의 그래픽 표현물로 전환시킨다. 그 다음 전환된 표현물들을 영상으로 생성하고 전시하는 과정을 거친다. 이후 정보공간에서 이루어지는 사용자의 방문행동에 의해 변화하는 관심도 또는 어떤 매개변수(parameter)에 따라 정보 통제를 실행한다. 예를 들면 이미지의 확대, 회전, 숨김 등이 그런 것들이다(Yang et al., 2003).

프로그램, 데이터베이스, 온라인 텍스트

등의 대형 정보구조체를 표현하는데 유용하게 적용될 수 있는 방법 중에 fisheye views가 있다(Furnas, 1986). 기본적으로 fisheye views는 세부사항(local detail)과 전반적 구조(global context)를 구별하여 대응하고자 하는 데에서 출발하였다. Furnas(1986)는 다음과 같은 관심도(Degree of Interest) 함수를 이용하여 여러 항목간의 거리를 표현하였다.

$$DOIfisheye(x, y) = API(x) - D(x,y) \dots\dots (1)$$

DOIfisheye 는 사용자가 현재 초점을 맞추고 있는 위치를 y라고 놓았을 때, x라는 점이 가지는 관심도의 크기를 나타내는 함수이다. API는 x의 a priori 중요도, D는 x, y의 거리를 의미한다.

Furnas(1986)는 인간이 사물이나 현상을 바라봄에 있어서 서로 다른 견지를 가지게 되는 현상을 다음과 같은 예를 들어 설명하였다. 심리학자들은 (경영, 마케팅)이라는 하나의 쌍과 (심리학, 정신의학)이라는 또 다른 쌍을 대할 때, (경영, 마케팅)의 쌍을 더 유사한 원소의 집합으로 여긴다. 심리학자들과는 달리 반대로 경영학자들은 (경영, 마케팅)의 쌍을 분명히 다른 원소의 집합으로 인식하고 있다. 이러한 일련의 예를 들면서, Furnas(1986)는 물고기눈(fisheye)처럼 가까이에 있는 것은 상세히 보이지만, 멀리 떨어진 곳에 있는 것은 덜 상세하게 보이는 이런 현상을 fisheye views라 지칭하고, 이의 여러 활용방안을 제시하였다.

April 1986		S	M	T	W	Th	F	S
Apr 6	6	7	8	9	10	11	12	
		*Prepare experiment terminal for expt/demo check f.e. tree pgm on ambassador terminal *Frank Halasz arrives in NE call about: dinner w/hin&no *J.A.B. running subjects 3 per day	*GWF runs subject 3-5pm in lab *DEPT WORK REVIEW Monna & Lee 10am 2q309 *Dinner w/ Frank & Mo? *J.A.B. running sub 2 today	*FRANK HA visitin *HALASZ T 10am Noteca 2L-12 *TKLGP Mt ->comme on s *J.A.B. r 3 per d	*GWF re (*MAL 11 (*Inf 1: *J.A 2: 3 T	*TOM vi *MAL 11 (*Dia 2: T	*MLS Fo a	
Apr 13	13	14	15	16	17	18	19	
	*FLY 9: P *Che Pa R	*BOSTON-CHI86 *DEMO REHEARSAL dont yet know where *Contact SYMBOLICS get denos ready *models workshop (place unkn.) *Register for conference	*BOSTON-CHI86 *GWF CHI86 talk 11am Salon F	*BOSTON-C *GWF DEMO 1:30-2p Regis	*BOS *Hil in H	*Don 2p N		
Apr 20	20	21	22	23	24	25	26	
		*PHIL BARNARD visiting *Bill Jones *DINNER	*Bill Jones Talk 11am *Phil Barnard visit *EZRothkopf talk		*JFP pi		*Dup *STD ca	
Apr 27	27	28	29	30	1	2	3	
		*To Texas	*MCC PTAC	*MCC PTAC	*MCC	*MCC C.	*NYC	

그림 1. Fisheye 달력(출처: Furnas, 1991)

예를 들자면, 그림 1.은 fisheye views를 적용하여 구성한 달력이다. 그림 1.과 같이 한 화면에 달력의 전반적 구조를 모두 보이게 하되, 오늘 또는 이번 주와 같이 관심있는 부분은 세부사항의 크기를 크게 표시되도록 하는 것이 fisheye views의 개념을 적용시킨 하나의 예이다. 이와 같이 화면표시와 인터페이스에 유용한 fisheye views 개념을 기초로 해서, 사용자의 관심도에 따라서 다른 크기로 항목이 제시되는 여러 방법들이 연구되었다. 예를 들자면 fisheye views 개념을 토대로 웹사이트의 많은 웹페이지와 콘텐츠들을 표시하는 방법이 고안되었다(Collaud et al., 1995; Fisher et al.,

1997; Heo and Hirtle, 2001; Turetken et al., 2004).

2.2 데이터의 인접성과 인접행렬

데이터의 인접성(adjacency)은 의사결정 공간(decision space)에서 다양한 관계를 이해하는 데에 유용하게 쓰일 수 있다. 품목 i 와 품목 j 가 동시구매되었는지, 또는 품목 i 의 구입이 품목 j 의 구매에 영향을 주었는지의 여부가 모두 데이터의 인접성이라고 할 수 있다. 데이터마이닝의 기법 중 연관규칙 분석(Berry and Linoff, 1997)이나, 또는 추천 시스템(Schafer et al., 2001) 및 데이터 시각화(Condon et al., 2002) 등이 데이터의 인접성을 이용한 몇 가지 예라고

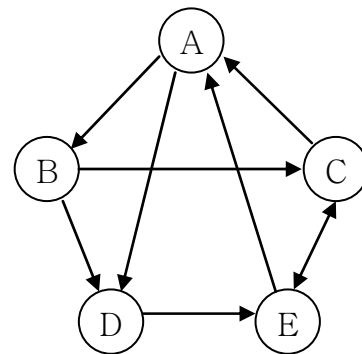
할 수 있다.

이러한 데이터의 인접성은 그래프 이론에서 출발하는데 유한개의 점과 선으로 구성된 도형을 그래프 또는 연결그래프(graph or connected graph)라고 하며, 사상(mapping)의 개념이나 도형의 위상적 성질을 명확히 알 수 있다. 이러한 그래프는 단위 정보를 표현하는 점과 각 점을 잇는 선으로 구성되어 있다. 또한 선에 방향성이 있는가에 따라서 유향 그래프(directed graph)와 무향 그래프(undirected graph)로 구분한다. 그 밖에 그래프 내에서 여러 점들의 연결과정을 경로(path)라고 하며, 시작점과 끝점이 연결된 경로는 특별히 순환(cycle)이라고 한다. 이를 행렬로 표현할 수 있는데, 연결그래프 내에서 선이 점을 공유하고 있으면 1, 공유하고 있지 않으면 0으로 나타낸 행렬을 인접행렬이라고 한다. 어떤 인접행렬 A는 n 개의 점을 가지는 $n * n$ 정방행렬이다. 인접행렬의 어떤 원소 $A_{ij} = 1$ 이면 두 점이 인접해 있다는 것이며, 또한 $A_{ij} = 0$ 이면 두 점은 인접해 있지 않은 것이다. 인터넷은 여러 통신망의 집합체이며, 각각의 통신망에 퍼져 있는 수많은 개체들이 서로 연결되어 있고, 또한 인터넷 사용자들은 어떤 순서와 방향을 가지면서 활동한다. 따라서, 인터넷 공간에 인접행렬의 개념을 이용할 수 있다(Błazewicz et al., 2005). 이러한 인접행렬에 기반하여 웹사이트를 구성하는 여러 단위들을 하나의 점으로 놓고 이에 방향성을 부여한 뒤, 각 사용자의 방문빈도를 측정할 수 있다. 예를 들어서 어떤 온라인 쇼핑몰에 총 5개의 상품(A, B, C, D, E)이 판매되고 있다고 가정한다.

사용자 1 = { A → B → D → E → C → A }
 사용자 2 = { A → B → C → E → A }
 사용자 3 = { A → D → E → A → B → D → E → C → A }

	A	B	C	D	E
A	0	3	0	1	0
B	0	0	1	2	0
C	2	0	0	0	1
D	0	0	0	0	3
E	2	0	2	0	0

(인접행렬)



(그래프)

그림 2. 그래프와 그 인접행렬의 예

이 때 3명의 사용자가 각각 그림 2.의 윗쪽 부분에서와 같이 순차적으로 이동하면서 쇼핑활동을 하였다면, 그 때의 그래프와 인접행렬은 그림 2.의 아랫부분과 같이 나타낼 수 있다. 그림 2.의 (그래프) 부분을 보면 유향성(directness)을 지니고 있음을 나타내기 위해 화살표로 방향성을 표시하였다. 또한 그림 2.의 (그래프)를 토대로 작성한 인접행렬이 그림 2.의 (인접행렬) 부분에 나타나 있다. 인접행렬은 0과 1의 값만을 가

하지만 여기에서는 방향성에 횡수를 더하였다. 방향성을 인식하기 위해 항상 행에서 열로 이동한다고 규정하기로 한다($i \rightarrow j$). 이 경우 A와 B를 놓고 보면, A에서 B로 3회 만큼 이동하였음을 알 수 있고, 반대로 B에서는 A로 이동한 횡수가 없다는 것을 알 수 있다. 따라서 방향 그래프는 보다 방향성이 있는 정보를 제공하고 있으며, 품목 간에 잦은 이동횡수가 나타나 있는 것, 즉 인접행렬의 어떤 원소의 값이 클수록 높은 인접성을 보인다고 할 수 있으며 이는 두 원소 간에 강한 연결(strong tie)이 존재한다는 것을 의미한다.

3. 시스템 구현

그래프와도 유사한 형태를 보이는 나무 구조는 수직적 체계(hierarchy)를 보여주기 때문에 직관적인 표현에 유용하지만, 수직 체계가 깊어질수록 공간을 많이 차지하므로 시각화에 부적절한 단점도 있다(Johnson and Shneiderman, 1991; Shneiderman, 1992). 컴퓨터에 저장된 파일 구조를 확인하기 위해 흔히 사용하는 윈도우 탐색기와 같은 것이 그 예가 될 수 있다.

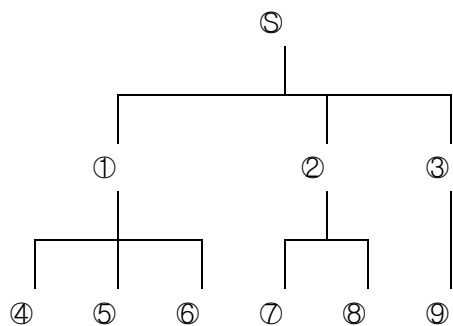


그림 3. 수직적 나무구조의 표현

그림 3.에는 간단한 수직적 나무 구조가 표현되어 있다. 만일 웹사이트를 방문한 어떤 사용자의 위치가 S에 있고 그 상태에서 세부적인 노드가 1 ~ 9 까지 있다고 가정한다. 그림 3.에서 세부 노드들의 가중치가 동일하다고 가정하면 첫 번째 수준의 노드인 1,2,3은 두 번째 수준의 노드를 각각 3개, 2개, 1개씩 차지하고 있기 때문에 한 화면에서 차지하는 영역의 크기비율이 각각 3:2:1로 나누어짐을 알 수 있다. 그림 3.의 수직적 나무구조를 참조해서 한 화면 내에 모든 최하위 노드인 두 번째 수준의 노드들을 구현한다면 최하위 노드의 수는 6개 이므로 화면의 비율을 각 노드당 약 17% 정도를 차지할 것이다.

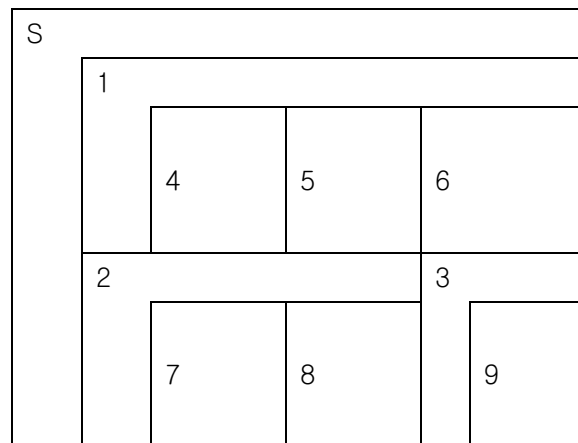


그림 4. 수직적 나무 구조의 모든 하위 노드들을 혼합하여 구현한 구조

전술했듯이 그림 4.는 사용자가 S의 위치에 있는 경우 두 번째 수준의 노드들을 같은 가중치를 가지고 있다는 것을 가정하는 상황에서 하나의 화면을 분할한 것이다. 그런데 모든 노드의 가중치가 항상 동일하지는 않고, 또한 이 가중치는 변화할 수 있다. 사용자가 어떠한 웹사이트 내의 많은 개체

들을 방문한 이력을 살펴보면 수직적 나무형 구조를 이루고 있음을 알 수 있고, 이예 더하여 개체 간의 연관성도 발생하는 것을 인식할 수 있다. 이러한 웹사이트내의 정보들이 이렇게 수직적 나무구조를 이루면서도 각 정보들 사이에는 데이터 연관성이 존재한다. 인터넷은 통신망의 집합이며 네트워크 구조의 표현이므로 현재 사용자의 위치가 ⑤인 경우 데이터의 인접성은 그림 5.와 같은 인접행렬로 나타낼 수 있다.

S	1	2	3	.	.	.	9
1	a1	a1	a1	.	.	.	a1
	1	2	3	.	.	.	9
2	a2	a2
	1	2
3	a2
	1
.
.
.
9	a9	a9
	1	9

그림 5. 그림 1.의 수직구조를 이루고 있는 노드들의 인접행렬

그림 5.에서 행에서 열 방향으로 움직인다고 규정하는 인접행렬의 성질에 의거해서 모든 노드의 수가 k개일 때, 상위 계층의 노드 중 i 노드의 가중치는 다음과 같이 정의된다.

$$\sum_{j=1}^k a_{ij} \dots\dots\dots (2)$$

예를 들면 1번 노드의 가중치는 다음과 같이 계산된다.

$$\sum_{j=1}^9 a_{1j} = a_{11} + a_{12} + \dots + a_{19} \dots(3)$$

따라서 상위 노드인 1,2,3의 가중치를 고려하여 각 노드가 화면에서 차지하는 비율은 다음과 같다.

node 1 : node 2 : node 3 =

$$\sum_{j=1}^9 a_{1j} \quad \sum_{j=1}^9 a_{2j} \quad \sum_{j=1}^9 a_{3j} \dots\dots(4)$$

한정된 화면에 하위 노드들을 모두 배치하는 것은 정보과잉을 야기시키기 때문에 가용공간을 보다 효율적으로 이용하려면 관심영역을 확대할 필요가 있고, 수많은 정보들을 적절하게 배치하기 위해 적정 가시도 수준(visibility threshold)을 유지하면서 화면을 구성해야 한다. 즉 상위 수준의 노드들이 가지는 가중치 합이 하위 수준 노드들의 가중치 합 보다 큰 경우 보다 상위 수준의 노드들로만 화면을 구성할 필요가 있다. 그림 1.의 예에서는 상위 노드 1,2,3의 가중치 합이 하위 노드 4,5,6,7,8,9의 가중치 합 보다 큰 경우가 이에 해당되며 다음과 같이 나타낼 수 있다.

$$\sum_{i=1}^3 \sum_{j=1}^9 a_{ij} \quad \rangle \quad \sum_{i=4}^9 \sum_{j=1}^9 a_{ij} \dots\dots(5)$$

수직적 나무 구조하에서 여러 계층이 나타날 수 있는데 하나의 계층은 군집을 이루는 개체들의 모음이라고 할 수 있다. 식 (5)의 경우 우변의 값이 크다면 그림 3.의 최

하위 계층을 이루는 4,5,6,7,8,9의 노드들로 한 화면을 구성하게 된다. 여기에서는 상위 계층이 하위 계층보다 가중치 합이 크다고 가정하고, 그 경우 상위 수준의 노드인 1,2,3을 대상으로 전체 확대화면을 구현하면 그림 6.과 같다.

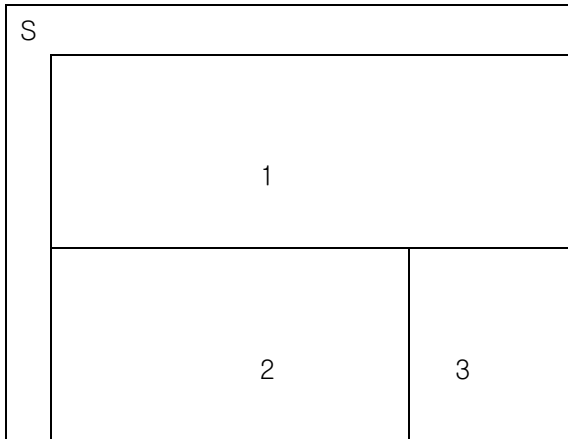


그림 6. 그림 1.의 상위 노드를 대상으로 최 확대하여 구현한 구조

그림 6.은 현재 사용자가 S의 위치에 있는 경우 상위 노드들을 대상으로 최대한의 확대화면을 제시하는 것이다. 이 때 노드 1만을 대상으로 노드 1이 가지고 있는 상세한 내용을 알고 싶은 경우 노드 1의 하위 노드들을 표현하여야 한다. 하위 수준 노드의 가중치는 상위 수준의 일부분을 차지함을 알 수 있다. 1번 노드의 경우 하위 노드가 4, 5, 6의 세 가지이다. 따라서 하위 노드 4, 5, 6이 차지하는 각각의 화면비율은 다음과 같다.

$$\text{node 4} : \text{node 5} : \text{node 6} = \frac{a_{14}}{\sum_{j=1}^9 a_{1j}} : \frac{a_{15}}{\sum_{j=1}^9 a_{1j}} : \frac{a_{16}}{\sum_{j=1}^9 a_{1j}} \dots\dots(5)$$

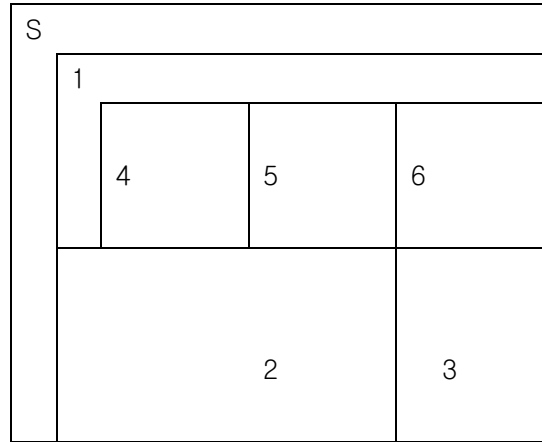


그림 7. 데이터 인접성에 기반한 fisheye views 확대를 구현한 구조

최종적으로 S의 위치에 있는 사용자에게 제시하게 되는 화면은 그림 7.과 같다. 이와 같은 방식으로 관심도가 높은 부분의 정보들만을 하나의 정보표현 공간에 모두 확대하여 제시할 수도 있고, 또한 상위 수준의 영역을 유지하면서 하위 수준으로 전개해 나가는 보다 fisheye views 관점의 확대를 통한 정보표현 방법도 구현할 수 있다.

4. 유용성 실험

본 연구에서 제시하는 시각화 방법의 효과를 검증하기 위하여 유용성 실험을 실행하였다. 실험의 목적은 본 연구에서 제안하는 데이터 인접성에 기반한 fisheye views를 통한 시각화 구현방법이 웹사이트 방문자의 정보탐색 및 정보획득에 도움을 주는 지 알아보려는 것이다. 시각화 기법의 효과는 주로 과업달성 여부로 측정하며(Keim, 2000; Turo and Johnson, 1992), 과업달성에 이르기까지의 시간과 그에 따른 만족도

를 성과지표로 측정된다(Turetken and Sharda, 2001; Yang et al, 2003). 이를 위해 가상의 온라인 쇼핑몰을 구축하였으며, 대학생과 대학원생들로 구성된 70명의 피험자들로 하여금 이 가상의 온라인 쇼핑몰을 방문하도록 하였다. 실험은 2단계에 걸쳐 시행되었는데 1단계에서는 기본적인 데이터 인접성에 관한 자료를 모으기 위하여 실행되었다.

70명의 피험자들로 하여금 실험을 위해 제작한 가상 온라인 쇼핑몰을 자유롭게 방문하여 탐색하도록 하였고, 이러한 방문 행동 데이터를 수집하여 데이터 인접성을 판단하기 위한 그래프 및 인접행렬로 전환하였다.

이를 토대로 fisheye views 기반의 시각화 기법이 구현된 웹페이지를 디자인하여 제시하는 2단계 실험을 진행하였다. 1단계 실험에서의 피험자 70명으로 하여금 해당 가상 쇼핑몰을 재방문하게 한 다음 역시 자유롭게 방문하여 탐색 및 정보획득과 같은 행동을 취하도록 하였다. 그 과업이 종료된 시점에서 완수되기까지의 시간과 만족도를 측정하였다. 2단계 실험에서는 비교검증을 위하여 총 3가지의 서로 다른 시각화 기법이 적용된 웹페이지를 각 피험자에게 제시하여 경험토록 한 뒤 그 때의 시간과 만족도를 측정하였다. 이 세 가지 시각화 기법은 각각 아무런 시각화 기법을 제시하지 않은 일반적인 상태(no technique), 필요한 부분만 최대 확대한 상태(full zoom), 그리고 본 연구에서 제안하는 데이터 인접성으로 구현한 fisheye views로 구현한 상태(fisheye views on data adjacency) 등이다. 구축한 화면의 예는 그림 8, 9, 10.에 제시되어 있다. 먼저 그림 8.은 본 가상 온라인

쇼핑몰에 접근하는 사용자가 보게 되는 일반적인 화면이다. “전자제품”과 같은 특정 항목을 선택한 경우 해당 항목에 포함되어 있는 일련의 제품군을 나열하고 있는 모습이다.

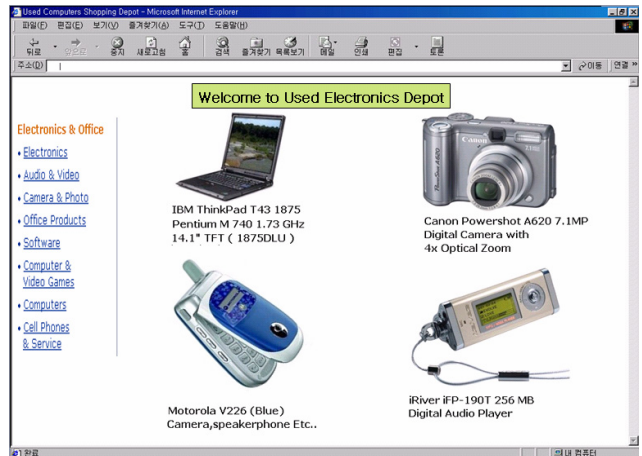


그림 8. TYPE I: 일반 웹페이지 화면(No technique)

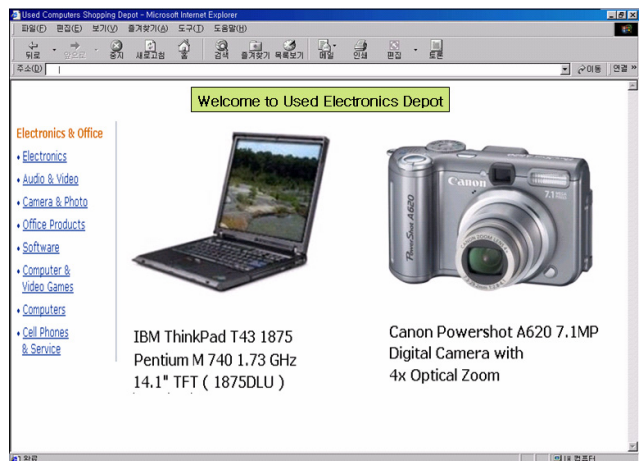


그림 9. TYPE II: 관심제품을 확대하여 표현한 웹페이지 화면(Full Zoom)

따라서 항목내에 있는 모든 제품들은 거의 대등한 크기의 화면비율을 점하며 표시된다. 이제 한정된 화면 내 공간을 몇 가지 제품이 본래의 크기 보다 확대된 상태로 표

현되어 있는 것이 그림 9.이다. 그림 9.에 나열되어 있는 제품들은 그림 8.의 개체들과 비교하여 볼 때 대부분 크기는 크고 나열된 제품 수는 적다. 이 결과는 상기 설명한 것처럼 데이터 인접성에 기반한 것으로써 수직적 나무 구조하에서 많은 연관성을 보이고 있는 개체들의 집합이라고도 할 수 있다. 그림 10.에서 구현된 화면이 바로 본 연구에서 제안하는 데이터 간의 연관성을 인접행렬 기반으로 해서 fisheye views 개념으로 시각화하여 표현한 것이다. 그림 10.에서 의미하는 바는 사용자가 제품을 검색 중에 자신의 위치를 잃어버리지 않도록 전반적인 구조를 제시하면서도, 현재 관심도가 높은 정보를 알려주고 있다는 점이다. 이러한 방식으로 사용자가 정보를 검색하고 분석하는 과정을 돕도록 시각화하여 설계된 화면을 구성할 수 있다.

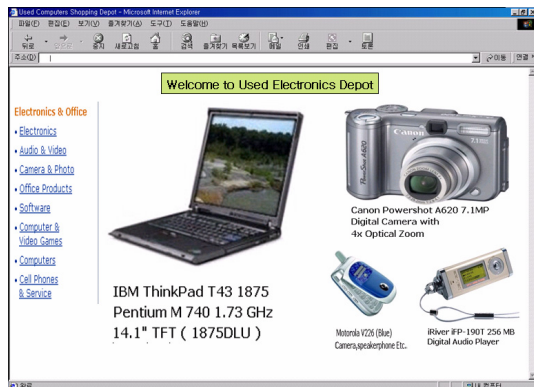


그림 10. TYPE III: Fisheye views 확대 웹페이지 화면

실제 실험에서 측정단위를 정하기 위해 시간은 초 단위로 설정하였으며, 사용자 만족도 부분은 “기대 이상의 방문경험” 그리고 “사전 신념과 선택대안과의 일관성”을 세부 측정항목으로 삼은 Brewer(2000), McKinney et al.(2002)의 연구를 참조하여

5점 척도로 측정하였다. 이러한 서로 다른 시각화 기법 간의 차이를 분석하기 위하여 일원분산분석(ANOVA)을 시각화 성과분석과 함께 실행하였다. 이 때의 분석결과는 표 1.에 나타나 있다.

표 1.에 나타난 실험 결과를 보면, 먼저 사용자 만족도 측면에서 관심도가 높은 부분을 확대하여 제시하는 시각화 방법이 만족도가 높음이 나타났고, 특히 데이터 인접성에 기반한 fisheye views 방식의 화면구성 방식이 높은 만족도를 사용자에게 부여하고 있음을 알 수 있다. 그리고 과업완수 시간의 요소에서도 이른바 확대한 정보를 제시하는 방법이 사용자에게 높은 시각화 효과를 줌으로써 정보검색 및 정보획득에 걸리는 시간이 단축됨을 알 수 있다. 또한 서로 다른 시각화 방법의 차이가 유의한지 알아보기 위해 실시한 분산분석(ANOVA) 결과 사용자 만족도나 과업완수 시간 등 모든 면에서 유의수준($\alpha = 0.05$) 내에서 충분히 유의하므로 제시한 시각화 방법이 높은 효과를 가지고 있음이 통계적으로 유의하다고 할 수 있다. 분석결과를 보다 자세히 확인하기 위하여 시각화 기법 간의 차이를 t-test 분석을 이용하여 살펴 보았다(표 2.). 표 2.의 결과를 보면 TYPE I 과 TYPE II의 차이는 상대적으로 덜 유의함을 알 수 있다. 그에 비해서 TYPE III의 시각화 효과는 다른 시각화 기법보다도 만족도나 시각 면에서 모두 충분히 유의함을 알 수 있다. 이런 결과를 볼 때 정보과잉 상황에서 필요한 정보만을 추출해서 보여주는 것이 필요하긴 하지만(TYPE II), 시간이 지나고 방문활동이 지속되면서 공간 내에서 방향상실이 가속되는 상황은 피할수 없게 된다는 점을 알 수 있다.

표 1. 시각화 기법 간 만족도와 시간에 대한 ANOVA 결과(N = 90)

	적용 기법	평균	표준편차	F-ratio	p-value
사용자 만족도 (5점 척도)	TYPE I (No technique)	3.11	0.43	24.362*	.000
	TYPE II (Full Zoom)	3.21	0.56		
	TYPE III (Fisheye views)	3.58	0.38		
과업완수 시간 (초 단위)	TYPE I (No technique)	587.6	63.7	6.062*	.002
	TYPE II (Full Zoom)	539.3	60.6		
	TYPE III (Fisheye views)	501.9	59.3		

* 유의수준 $\alpha = 0.05$ 에서 통계적으로 유의함.

표 2. 두 가지 시각화 기법간 t-test 결과

				p-value
TYPE I (no technique)	vs	TYPE II (full zoom)	만족도	0.054
			시간	0.047
TYPE I (no technique)	vs	TYPE III (fisheye views)	만족도	0.005
			시간	0.000
TYPE II (full zoom)	vs	TYPE III (fisheye views)	만족도	0.009
			시간	0.001

따라서 본 연구에서 제안하는 데이터 인접성에 기반한 fisheye views를 이용하여 시각화한 웹페이지가 사용자에게 유용함을 알 수 있다.

5. 시사점 및 결론

현재의 인터넷 환경에서는 지나치게 많은 정보와 다양한 매체의 제공으로 인한 이른바 인터넷 사용자들의 “초공간에서의 방향 상실(lost in space)” 현상이 심화되고 있다. 따라서 지나친 정보과잉(information overload)을 해소하면서도, 사용자가 필요로 하는 정보를 효과적이면서 효율적으로 제시할 필요가 있다. 한정된 공간 내에 여

러 결과물을 표시하여 정보공간을 보다 유용하게 활용하기 시각화 기법은 컴퓨터 화면이라는 한정된 공간을 다루는 데에도 필요한 방법이다.

본 연구에서는 여러 시각화 기법 중에서도 전체적 맥락(context)과 상세 내용 보기(details)를 통합시켜 제시하는데 유용하게 쓰이는 fisheye views를 웹페이지 시각화에 적용하였다. 또한 fisheye views 개념의 중심적 내용인 관심도의 차이를 규정하기 위해 그래프 이론에 기반을 두고 있는 데이터 인접성을 이용하였다. 그래프 이론은 다양한 조합적 최적화에 기인한 많은 문제의 분석에 유용하고, 특히 인터넷과 같이 많은 양의 데이터가 존재하는 정보공간에서의 문제해결에 큰 이점이 존재한다. 이러한 제안된 시각화 기법의 유용성을 검증하기 위하여 가상 인터넷 쇼핑몰을 구축하였고, 여기에 방문하는 사용자들에게 서로 다른 시각화 기법이 적용된 웹페이지를 제시함으로써 그 효용도를 검증하는 실험을 실행하였다. 서로 다른 시각화 기법으로는 어떤 시각화 기법도 제시하지 않은 일반적 형태의 웹페이지와 특정 정보의 최대 확대 화면이 구현된 웹페이지, 그리고 본 연구에서 제시한 데이터 인접성에 기반한 fisheye view 등을 피실험자에게 경험토록 하였으며, 과업완수에 이르기 까지 걸리는 시간과 과업완수후의 만족도를 조사하였다. 결과적으로 데이터 인접성과 fisheye views를 결합한 본 연구에서의 웹페이지 시각화 기법이 적은 과업완수 시간을 들이고도 만족도는 높은 결과를 나타내어 유용함을 입증하였다.

본 연구에서는 웹을 구축하고 설계하는 한 가지 유용한 방법론을 제시하였다. 향후 연구에서는 본 연구에서 처리하지 못한 점

들을 보완할 필요가 있다. 우선 데이터 인접성에 의해 주목받는 부분에 대한 속성을 비교분석해 볼 필요가 있을 것이다. 밀접한 관계를 맺고 있는 개체 사이에 발생하는 특이한 속성 등을 발견함으로써 연관성 이상의 정보를 획득할 수도 있기 때문이다. 또한 실험을 위한 연구대상 웹사이트로서 인터넷 쇼핑몰을 상정하였는데 이러한 상업적 웹사이트가 아닌 검색엔진과 같은 정보제시에 보다 주의를 두는 웹사이트에도 적용해 볼 수 있다. 그리고 시각화 효과를 평가함에 있어서 시간과 만족도 뿐만 아니라 사용자의 순간적 반응성 등 보다 세밀하고 유용한 측정도구를 적용할 수 있을 것이다.

참고문헌

- [1] Becker, R.A., Cleveland, W.S. and Martin, R.D., "Trellis graphics displays: a multi dimensional data visualization tool for data mining," Proceedings of 3rd Annual Conference on Knowledge Discovery in Databases, Newport Beach, CA, August 1997.
- [2] Berry, M. and Linoff, G., Data Mining Techniques, Wiley, New York, 1997.
- [3] Błażewicz, J., Pesch, E. and Sterna, M., "A novel representation of graph structures in web mining and data analysis," Omega, Vol. 33, No. 1, 2005, pp. 65-71.
- [4] Brewer, M., "Measuring Customer Intimacy," Valoris Abram Hawkes, 2000, pp. 1-6.
- [5] Cleveland, W.S., Visualizing Data, AT&T Bell Laboratories: Murray Hill, NY, 1993.
- [6] Collaud, G., Dill, J., Jones, C.V. and Tan, P., "A Distorted View Approach to Assisting Web Navigation," in Proceedings of Workshop on New Paradigms in Information Visualization and Manipulation (in conjunction with 4th Int. Conf. on Information and Knowledge Management, Baltimore, December), 1995, pp. 95-100.
- [7] Condon, E., Golden, B., Lele, S., Raghavan, S. and Wasil, E., "A visualization model based on adjacency data," Decision Support Systems, Vol. 33, No. 4, 2002, pp. 349-362.
- [8] Fisher, B., Agelidis, M., Dill, J., Tan, P., Collaud, G. and Jones, C., "CZWeb: Fish-eye Views for visualizing the World-Wide Web," In Proceedings of the 7th International Conference on Human-Computer Interaction, 1997, pp. 719-722.
- [9] Furnas, G.W., "Generalized Fisheye Views," Published in Human Factors in Computing Systems CHI '86 Conference Proceedings, 1986, pp. 16-23.
- [10] Furnas, G. W., "The Fisheye Calendar System," Bellcore TM-ARH-020558, 1991, pp. 1-9.
- [11] Heo, Misook and Hirtle, Stephen C., "An empirical comparison of visualization tools to assist information retrieval on the Web," Journal of the American Society for Information Science and Technology, Vol. 52, No. 8, 2001, pp. 666-675.
- [12] Johnson, B. and Shneiderman, B., "Treemaps: a space filling approach to the visualization of hierarchical information structures," Proceedings of the Second International IEEE Visualization Conference, San Diego, Oct. 1991, pp. 284-291.
- [13] Keim, Daniel A., "Visual exploration of large date sets," Communications of the ACM, Vol. 44, No. 8, 2001, pp. 39-44.
- [14] Lamping, J. and Rao, R., "Visualizing large trees using the hyperbolic browser," Proceedings of the Conference Companion on Human Factors and Computing Systems, Vancouver, British Columbia, 1996.
- [15] McKinney, V., Yoon, K. and Zahedi, F.M., "The measurement of web-customer

- satisfaction: an expectation and disconfirmation approach,” *Information Systems Research*, Vol. 13, No. 3, 2002, pp. 296–315.
- [16] Muchaluat, Débora C., Rodrigues, Rogério F. and Soares, Luiz Fernando G., “WWW Fisheye-View Graphical Browser,” In *Proceedings of the 5th International Conference on Multimedia Modeling*, Lausanne, 1998, pp. 80–89.
- [17] Rivlin, E., Botafogo, R.A. and Shneiderman, B., “Navigating in hyperspace: designing a structure-based toolbox,” *Communications of the ACM*, Vol. 37, No. 2, 1994, pp. 87–96.
- [18] Schafer, J., Konstan, J. and Riedl, J., “E-commerce recommendation applications,” *Data Mining and Knowledge Discovery*, Vol. 5, No. 1&2, 2001, pp. 115–153.
- [19] Shneiderman, B., “Tree visualization with treemaps: a 2D space-filing approach,” *ACM Transactions on Graphics*, Vol. 11, No. 1, 1992, pp. 92–99.
- [20] Tufte, E.R., *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT, 2001.
- [21] Turetken, Ozgur and Sharda, Ramesh, “Visualization support for managing information overload in the web environment,” *Proceedings of the 22nd International Conference on Information Systems (ICIS)*, December 16–19, 2001, New Orleans, LA, USA.
- [22] Turetken, Ozgur, Schuff, David, Sharda, Ramesh, and Ow, Terence T., “Supporting systems analysis and design through fisheye views,” *Communications of the ACM*, Vol. 47, No. 9, 2004, pp. 72–77.
- [23] Turo, David and Johnson, Brian, “Improving the visualization of hierarchies with treemaps: design issues and experimentation,” *Proceedings of the IEEE International Conference on Visualization*, Los Alamitos, CA, USA, 1992, pp.124– 131.
- [24] Yang, Christopher C., Chen, Hsinchun and Hong, Kay, “Visualization of large category map for Internet browsing,” *Decision Support Systems*, Vol. 35, No. 1, 2003, pp. 89–102.